

第十一章-HEVC/H.265 的 HDR 编码方案

前文我们对 HEVC 的 HDR 编码优化技术做了介绍,侧重编码性能的提升。本章主要阐述 HEVC 中 HDR/WCG 相关的整体编码方案,包括不同应用场景下的 HEVC 扩展编码技术。

11.1 背景

HDR 信号一般意味着使用更多比特,一般的 HDR 信号倾向于使用 10 比特。这代表着更多的量化阶数,不同阶之间更平滑的过渡,但这也意味着更高的带宽和存储需求。未压缩的单个 HDR 像素需要 12 个字节的内存来存储 RGB 值的三个单精度浮点数。未压缩的 4K UHD 分辨率(3840×2160 像素)的单个 HDR 帧需要大约 94.92M 字节的存储,并且每秒 30 帧的一分钟数据需要 166G 字节。如果没有压缩,会给存储和传输带来巨大的压力。虽然地面宽带的带宽经常增加,但即使这样也不足以在未来播放任何未压缩的 HDR 电影,因此对数字信号进行压缩是进一步的需求。

HDR 视频的标准中,最常用的是由 BDA(Blu-ray Disc Association)支持的 HDR10 标准,这是最早支持 HDR 视频的几种标准之一。HDR10 标准包含的内容有:使用 PQ 电光转换函数和 BT.2020 标准中的色度空间,10 比特的非恒定亮度 Y'CbCr 空间,4:2:0 的色度抽样方法,压缩方法采用 HEVC Main 10 配置文件。在实际应用时,HDR10 中还应该包括一些常用的元数据,比如 Mastering Display Colour Information 或者 Content Light Level Information 等。

2013 年 7 月,MPEG 开始针对 HDR 和 WCG 内容分发的需求,展开探索性的实验,打破了视频压缩和处理链的各部分,转换函数被分离研究,不同的颜色表示格式测试各种抽样方法,并且做了一些主观视频质量测试方法和客观视频质量测试指标。MPEG 最初的重点是高质量的应用程序,但后来扩大到包括广播,流媒体和存储应用程序。随后,MPEG 于 2015 年 2 月发布了证据征集(CfE),并于 2015 年 6 月报告了对技术响应的评估。CfE 的目标是评估当前 HEVC 标准的 HDR 和 WCG 视频的压缩性能,并确定是否应考虑其他有前景的技术以用于 HEVC 的未来扩展。

11.2 HEVC 编码对 HDR/WCG 的支持

在 H.265/HEVC 针对 HDR 和 WCG 的 CfE 中,得到了两种类型的解决方案:一种是单层的 HDR 编码方案,包括元数据的使用,但并不需要对落后的 SDR 设备的向后兼容性;第二种明确的关注对 SDR 设备的兼容性,使用单层或多层的编码方案。这里指明的对 SDR 的向后兼容性不仅包括对现有编码器,也就是 HEVC Main 10 的兼容,也包括了对现有的 SDR 显示器的兼容。对于单层 HDR 编码,作为对比的参考技术是 HEVC Main 10;对于多层 HDR 编码,参考技术是 SHVC Scalable Main 10。HDR 和 WCG 技术需要的位深至少为 10bit,在电视制造业中,HEVC Main 10 已经广泛应用在 UHD 视频中。因此,HEVC Main 10 已经作为由 CfE 提出的技术中的单层案例的方法和可分级案例中的基础层。

第一和第二版本的 HEVC 标准已经包含了几种可应用于 HDR 和 WCG 视频的工具,包括支持 HDR 信号传输和 BT.2020 色域的 HDR 视频信号格式、用于支持 HDR 信号在接收端转换显示的补充增强信号(SEI)信息的定义和支持更高位深视频编码的应用。除此之外,

HEVC 的可分级扩展 (SHVC) 还包括色域可扩展性 (CGS) 和位深可扩展性 (BDS) 工具来解决针对 SDR BT.709 的向后兼容性。表一概括了 HEVC 编码器中为 HDR 和 WCG 提供的一些工具，而这些工具将在接下来的几个部分分别介绍。

表 1 HEVC 中 HDR/WCG 工具

Metadata / Tools	Purpose
VUI	Provides information related to the coded video signal “container”: color primaries, transfer function, and color space to map video sample code levels to display intensities. Besides the traditional ITU-R BT.709 combination required by DVD, ATSC, DVB, and Blu-ray 1.0, some camera log transfers and the new transfer function SMPTE ST 2084 [33], among others can be indicated. BT.2020 primaries were added to AVC and HEVC in 2013, followed by XYZ in 2014.
tone mapping information SEI message	Post-decode downloadable tone mapping curve applied to RGB primary signal. Includes multiple methods to transmit one or more curves compactly within the bitstream, where each curve targets a different mapping scenario. Note: originally developed for AVC in 2005.
mastering display color volume SEI message	Signals brightness range, primaries, and white point of the display monitor used during the grading of video content (SMPTE ST 2086 [34]).
color remapping information SEI message	Post-decode downloadable intensity map LUTs and 3x3 matrix for bit depth (dynamic range) and color space conversion.
knee function information SEI message	Post-decode downloadable tone mapping curve applied to linear (brightness) signal on all three RGB channels.
CGS/BDS	Defines color mapping between base layer and SHVC enhancement layer (e.g. mapping from BT.709 base layer to BT.2020 enhancement layer), for bit-depth and color gamut scalability.

11.2.1 VUI

VUI (Video Usability Information) 是在 H.264/AVC 和 H.265/HEVC 序列参数集中携带的元数据元素的集合，它们共同描述了实际视频信号如何在样本流内存储和映射。术语“视频信号容器”起源于早期的电视和电影工程，早于文件容器概念。VUI 传达视频信号的含义，特别是整数视频采样值如何最终转换为线性显示强度，通常表示为 0 到 1.0 的浮点数信号乘以传递函数的强度范围，例如 SDR 视频的 100nits 和 HDR 视频的 10000nits。普通数字电视信号是经过量化的一系列样本，它们拥有各自的以下属性：比特深度、非线性 (Gamma) 映射和导数颜色空间，以及附加属性比如色度亚采样方式。HEVC 视频编码的研究者通常更

多的关注视频信号容器的属性，而不是视频编解码器及其参数，编解码器通常被当作“黑箱”来使用，它通过临时压缩，变为用于传输的视频信号，然后由解码器来解码和重新渲染为原始视频信号的表示，最终保持视频编码之前信号容器的原始含义。诸如格式转换和显示处理之类的解码后处理与 VUI 容器属性高度相关，以优化视频信号在目标显示器上的渲染。同样，VUI 容器的属性显著影响了视频信号格式转换的效果，因此为了得到内容的预期效果，对 VUI 中元数据的正确设置显得至关重要。图 1 给出了 VUI 和 SEI 元素在渲染过程即后处理过程中可能有应用的地方。接下来将介绍关键 VUI 比特流元素，其规范中的名称用斜体显示。

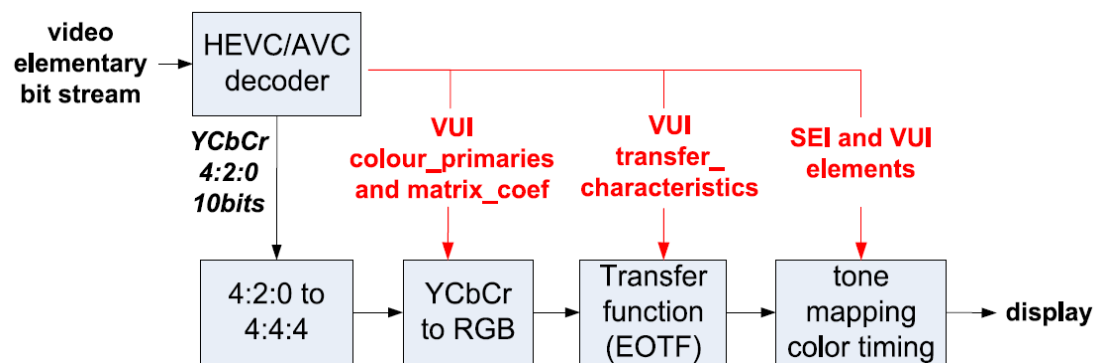


图 1 VUI 和 SEI 信息在渲染过程的应用

原色(colour_primaries)

视频中的原色可以由实际的 CIE 的 xy 坐标表示为 0-1 的数值，这个属性基本上建立了视频信号包含的色域的边界。关于 WCG，HEVC 在 VUI 中包含两个新条目，以支持 BT.2020 和 XYZ 颜色原色。如由 CIE xy 坐标规定的主要三角形轨迹所示，该区域可以在 colour_primaries 的指定值之间自由选取，并且 BT.2020 拥有两倍于 BT.709 的色域范围。

传递函数 (transfer_characteristics)

传递函数建立了样本量化值（0 到 $2^{\text{BitDepth}} - 1$ 之间的整数）与亮度显示器上的亮度（以 nit 为单位）的映射关系。VUI 指示符 *transfer_characteristics* 通常对应于参考的 SMPTE 或者 ITU-R 标准的特定方程，其被应用于将原始视频线性信号转换为非线性信号，比如传统的用于 SDR BT.709 的 Gamma 0.45 函数。大多数的传递函数将线性信号转换成更紧凑的非线性信号表示，然后用编码器进行压缩。线性信号一般有均匀的量化步长，因此占用的比特数比较多，不适用于低码率为主导的信号传输，例如，BT.709 中的逆 Gamma 传递函数将非线性 8 位信号扩展到 0.1 到 100 尼特范围内大约 11 位的线性光信号，这也从侧面显示了传递函数带来的码率节省。

新的 PQ 传递函数比通常使用的 Gamma 传递函数更适用于 HDR 视频信号，并且最近被添加到了 VUI 的传输特性表中。PQ 是 Barten 等人针对人类视觉模型对对比度敏感的特性来设计的，其被收录在 ST.2084 规范中。

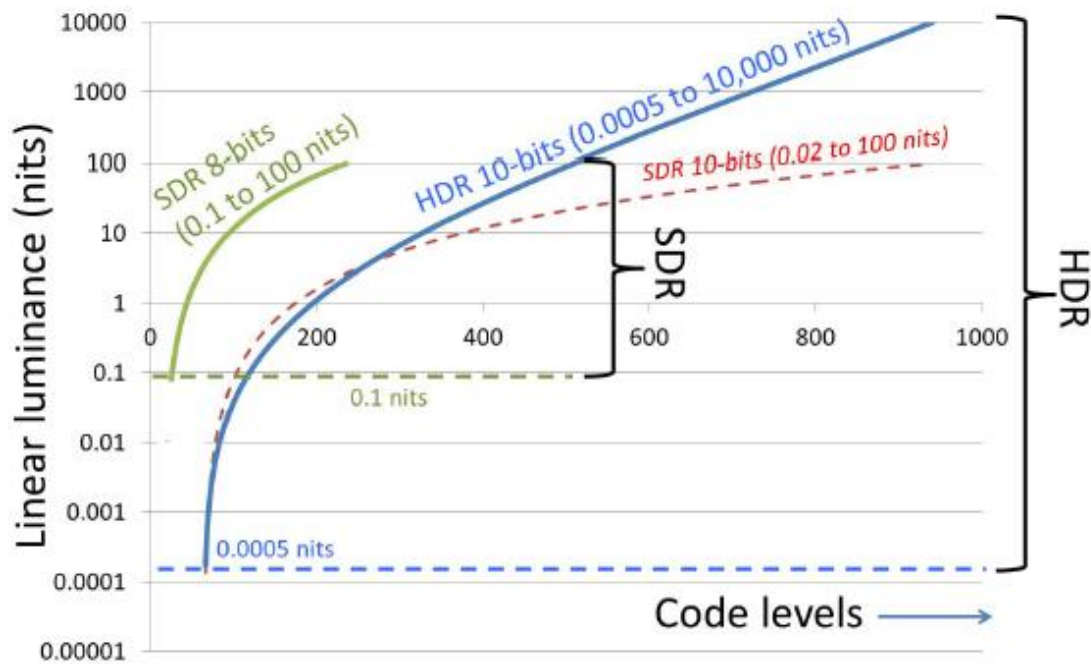


图 2 传递函数对比

PQ 传递函数与其他各种传递函数的对比如图 2 所示，该图的纵轴是亮度级别取 \log_{10} 的数值，单位是尼特，横轴是视频的整数样本的编码级别。作为对比，图中的绿色曲线代表了传统的用于 8 比特 SDR 视频的传递函数，蓝色曲线是用于 10 比特 HDR 视频的 PQ 传递函数。如图 5 所示，8 比特 SDR 仅仅覆盖了 0.1-100 尼特的范围，而 10 比特 HDR 仅仅通过额外的两个比特，覆盖了 0.00005 ~ 10000 尼特的范围。对于 BT.709 SDR 中能使用 8 比特（16-235）表示的亮度范围 0~100nits，PQ 使用了 9 个比特（120~510），另外一个额外位有效的提供了低于 0.1nits 的较暗处细节，而这些细节映射到 64~119 的量化值范围。而亮度大于传统 SDR 信号限制的 100nits 的部分被映射到 511~940 的量化值范围。需要注意的是，与专业的 10 比特 SDR BT.1886/BT.709 信号容器相比，对于 0.1-10nits 的信号，PQ 曲线密度相对较大，但这是一个权衡，对于 10~100nits 的信号，PQ 传递函数的曲线更加稀疏。

颜色空间(Matrix Coeffs)

非线性/近似感知统一的编码视频信号的实际域，从由上述传递函数和颜色原色的组合指示的非线性/感知统一的主视频信号导出，最常见的衍生色彩空间类似 Y'CbCr。除此之外，AVC 和 HEVC 认为编码的视频信号是原始信号，例如 4:4:4 RGB 或 XYZ，原始信号的编码在专业应用中是很常见的，例如相机主机或存档视频的夹层压缩。为了能够给出像 Y'CbCr 空间和 4:2:0 的采样比这样更加能量紧凑的色差空间同工的一些比特率上的优化，HEVC 第二个版本添加了交叉分量预测（Cross-component prediction, CCP）工具，基本思路是在解码重建的循环中，将原始信号分解为类似于 Y'CbCr 这样的不相关的几个分量，并用于预测。因此，CCP 是编解码器内部对 AVC 或 HEVC 编码之前应用于主信号的通常外部色差导出的内部替代方案，然后在 AVC 或 HEVC 解码之后通过逆操作处理得到原始信号的近似。

量化范围(video_full_range_flag)

量化范围定义了非线性编码的视频信号范围是否利用了从 0 到 $2^{BitDepth}-1$ 内所有可用的量化代码级别, 或者被缩放到某个子集用以提供足够的用于信号负脉冲的空间和用于信号过冲的空间, 用于存储通过从滤波器或编解码器量化产生的噪声。消费市场中的视频应用通常需要的是非全范围模式, 例如, 10 比特下常常将 0 到 1 之间的实数非线性信号值映射到量化值为 64-940 的亮度范围和 64-960 的色度范围。对比之下, 传统的 8 比特视频信号的亮度量化范围是 16~235, 色度是 16~240。

像所有 VUI 元素一样, *video_full_range_flag* 是在 HEVC 和 AVC 标准范围内的信息元数据, 因此不影响任何内部解码操作, 解码器仅需要将完整的动态范围内存储的输出视频和图像剪辑到内部解码图像缓冲区, 用作后续解码图像的运动预测参考。解码器不会通过语法元素 *video_full_range_flag*, 将视频映射到包含的任何较窄的范围, 而是将该 VUI 元素传递到最后用于显示器的处理。

在 MPEG HDR 和 WCG 特设组进行的实验中, 以及 Ultra HD Blu-ray 项目的多个调查中显示, 10 比特 PQ BT.2020 的 HDR、WCG 信号的编码比特率是与用于 SDR 编码比特率相同的数量级, 以 PSNR 为评价标准的 HDR 和 SDR 保真度也很接近。在测试序列中, HDR 技术主要体现在艺术效果中, 例如镜面高光、霓虹灯和分散在黑暗场景中的高对比度光和火焰, 实际场景的平均亮度级别被设置为与 SDR 视频中的平均亮度接近。除此之外, AVC 和 HEVC 在具有精细细节的镜面高光和高对比度边缘形状会使用更精细的四叉树划分 (比如使用 4x4 块)。与进行传统变换量化对比, PQ 方法进行更精细的划分, 有更少的残差和边信息, 因此带来的性能损失 (开销) 很少。

11.2.2 SEI 信息

SEI 是 Supplemental Enhancement Information 的缩写, 是用来保证 SDR 与 HDR 之间的互用性的机制的方法, 用于实现在接收器或者播放器中实现直接显示或转换显示的过程。色调映射是将 HDR 信号转换为较低动态范围的过程, 而反色调映射是它的逆过程, 旨在增强输入信号的动态范围和色域。这些过程通常由伴随着主视频流的元数据驱动, 在创建内容和分级时生成。HEVC 标准包含几种 SEI 信息, 可用于 SDR 和 HDR 之间和正常色域与宽色域之间的正向和逆向转换。

HEVC 标准的附录 D 中给出的 Mastering Display Color Volume SEI Message 对应了第一类型的指定表征内容参数的元数据, 这些参数是内容描述性的, 与解码信号后的具体处理无关。这种元数据表示了内容生成时监视器的实际色域和动态范围, 色域和亮度动态范围的乘积空间被称为“颜色容量”。通常情况下, 彩色视频的摄影师进行视频调整的创作端的母带监视器旨在代表接收端的现实目标, 在实践中, 母带制作显示器的颜色容量是编码的视频信号容器颜色容量的子集。例如, 如果分级监视器仅呈现 BT.709 色域, 但是信号容器指定了更大的 BT.2020 色域, 那么掌握显示颜色容量 SEI 信息中的参数就将指示 BT.709 空间的属性, 也就是 BT.709 的原色 xy 坐标和白点, SEI 信息还包含了显示器的最大最小亮度。知道母带监视器的子集范围会帮助到优化编码视频在显示设备上的显示效果, 这可能与母带制作中使用的不同。另外, 在接收端测量编码视频流实际使用的颜色容量也是可行的, 但这种分析会导致额外的延迟, 潜在地需要从开始到最后完整地传递内容以建立内容的可信赖的配置文件。

第二种元数据可以更具体的用于自适应过程, 这些元数据嵌入了控制色调映射和反色调映射的参数。他们通常针对特定的后处理过程, 例如通过查找表组合的信号转换和经由 3×3 矩阵的颜色空间转换。与上述的描述性元数据相反, 这种元数据可以认为是处理控制和参数的元数据, 例如图 3 中的元数据就可以被用于保证 HDR 转换得到的 SDR 信号符合本来的艺术层面的意义。这一类别的元数据, HEVC 规定了三种 SEI 信息: tone mapping 信息、颜色重映射信息 (CRI) 和 knee function 信息。

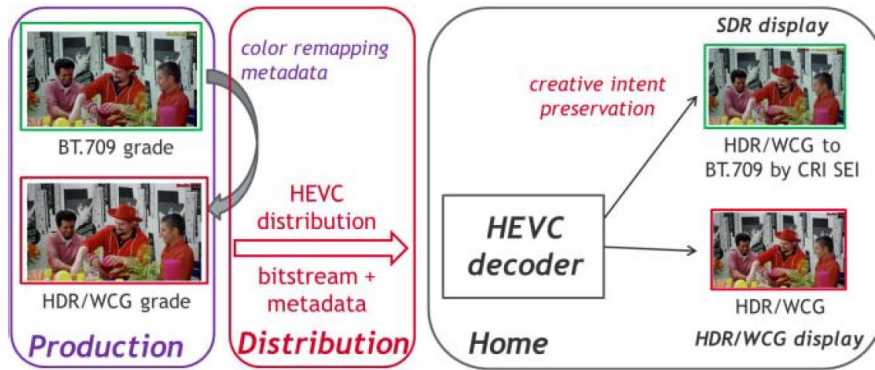


图 3 HDR 向 SDR 转换时 CRI 的应用

Tone mapping 的 SEI 信息被设计为携带编码视频序列内的一个或多个色调映射曲线, 色调映射曲线用于在显示器上显示解码图像数据, 该表达式通过导出在 SEI 信息中转录的一维查找表并将得到的表应用于解码信号的 RGB 分量来进行算术实现。为了提高编码效率, 在系统中的一维查找表给出了四种方式: 线性模型、S 形模型、分段线性模型和显式查找表模型。选择前两种的方法可以简化通用色调映射行为特征的信令, 后两种方法允许更多的任意映射, 但是具有更高的比特率上的潜力。在系统的设计中, 重要的是要注意通过色调映射信息 SEI 消息标识符标签可以将多于一个色调映射与视频编码序列相关联。这使得系统可以同时支持多个动态范围目标, 包括具有比解码的视频数据更大或者更小的动态范围的目标。该信息对应 *tone_map_id* 参数。该参数使得能够存在和维护多个色调映射信息 SEI 消息。而且, 应用程序可以使用该参数来识别所包括的色调映射。用于识别色调映射的目的第二种机制是对存在于比特流中的每个色调映射的编码数据位深和目标位深的信令。编码数据比特深度指示 SEI 消息所针对的编码视频序列的比特深度。当该值与比特流中的数据的比特深度相同时, 则可以将特定色调图视为使用的候选。或者, 当编码数据比特深度值与比特流的比特深度不匹配时, 则色调映射图可以被假设为 (一种情况是该序列经过了转码或在传输系统内已经进行了色调映射操作)。类似地, 色调映射图可以与不同的目标比特深度相关联。该信息指示接收设备或显示器的位深度要求。

颜色重映射 (CRI) SEI 信息是用来保证转换的 HDR 和 WCG 与 SDR 信号内容不失真, 适用于在生产、后期制作方面制作两个分级版本的内容时。CRI 元数据源自这两个分级版本, 并传送到解码器, 那么在后处理步骤中可以确保符合艺术意图的转换。重映射模型适用于信号的三种颜色分量信号, 例如 RGB 或 Y'CbCr。初始模型基于应用于图像的三个颜色分量样本的三维查找表。后来过程被简化, 但是没有明显处理质量上的影响, 以适应当前消费电子设备显示能力上的限制。最终版本基于三个一维分段线性输入查找表 (LUTin0, LUTin1, LUTin2), 随后是一个能够进行颜色空间转换的 3×3 矩阵, 紧接着三个一维分段线性输入查找表 (LUTout0, LUTout1, LUTout2)。以下等式说明了该模型在颜色样本 RGB 值中的应用:

$$\begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} LUT_{in0}[R_{in}] \\ LUT_{in1}[G_{in}] \\ LUT_{in2}[B_{in}] \end{bmatrix}$$

$$\begin{bmatrix} R_2 \\ G_2 \\ B_2 \end{bmatrix} = M_{3 \times 3} \begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix}$$

$$\begin{bmatrix} R_{out} \\ G_{out} \\ B_{out} \end{bmatrix} = \begin{bmatrix} LUT_{out0}[R_2] \\ LUT_{out1}[G_2] \\ LUT_{out2}[B_2] \end{bmatrix}$$

这三个步骤的每一步都是非强制的，例如在只有一个输入信号的传递函数对每个颜色分量的应用中，只有第一个一维查找表可以适用。由于输入和输出的位深可以不相同，因此该模型可以支持 SDR 与 HDR 的转换和不同色域之间的转换。

Knee function SEI 信息可以被认为是上面 Tone mapping SEI 信息中第三种可分段线性函数的简化版本，分段点在此定义为拐点。例如图 4 中给出的简单模型，该模型只包含三个拐点，说明了一个典型的 S 形曲线是如何通过拐点函数设计，来实现由一个 1000nits HDR 输入信号映射到预期为 100nits SDR 信号的输出信号。HEVC 提供的 SEI 信息中可以支持最多 999 个拐点的拐点方程。

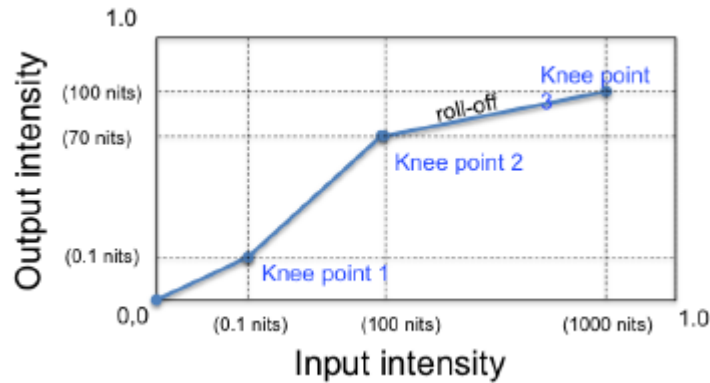


图 4 S 形 knee function 示意图

11.2.3 SHVC 中的编码工具——CGS、BDS

SHVC 是 HEVC 编码的可分级扩展功能，为了保持 SHVC 的低执行成本，没有引入新的块级编码工具。为了减少层间冗余，在编码增强层处采用上采样的基本层来重构图像作为层间参考图像。HEVC 的设计通过使用分层时间预测结构的时间子分层特征来支持时间可扩展性，SHVC 提供了额外的可扩展功能，包括位深、空域、质量、隔行扫描到逐行扫描、色域和标准（例如具有 HEVC 增强层的 AVC 基础层）的可扩展性，允许支持具有 SDR 后向兼容性的 HDR 和 WCG 信号。

色域可扩展性 (CGS) 表示基本层和增强层各自具有不同色域的情况下可扩展使用的例子。例如，DVB Phase II 定义了一种具有 8 比特 BT.709 色域的 HD 内容作为基础层的场景，

以及增强层中 10 比特 BT.2020 色域的 UHD 内容。显然，在两层之间不同色域特性的直接层间预测无法在 CGS 场景中工作，因此 SHVC 开发了 CGS 的特殊编码工具，为 SHVC 发布了两个配置文件：Scalable Main profile 和 Scalable Main 10 profile。其中 Scalable Main 10 profile 支持 CSG 相关编码工具。

CGS 的一般过程如下：首先，将配置的基础层重建图像映射到增强层的色域；然后，将映射图像（按需使用上采样）用作当前增强层图像的解码时的层间参考图像，就其本身而论，映射过程就是色调映射，实际上是作为 SHVC 中 CGS 的编码工具引入的唯一新模块。在 CGS 编码工具开发初期，对采样过程和色调映射过程的顺序进行了深入研究，以找出编码性能与编码复杂度之间的最佳权衡。SHVC 采用的阶段顺序在上采样之前进行色调映射，可以大大降低色调映射的计算复杂度，编码性能上只有不到 1% 的损失。

在颜色映射算法设计中有三个主要的元素：基于不对称三维查找表的 CGS、非均匀色度内容分割和亮度和色度分量之间的相位调整。与不使用 CGS 工具的实验对比，这三个元素的应用在不同的测试方案下，可以有 10% 到 15% 的 BD-rate 上的优化。

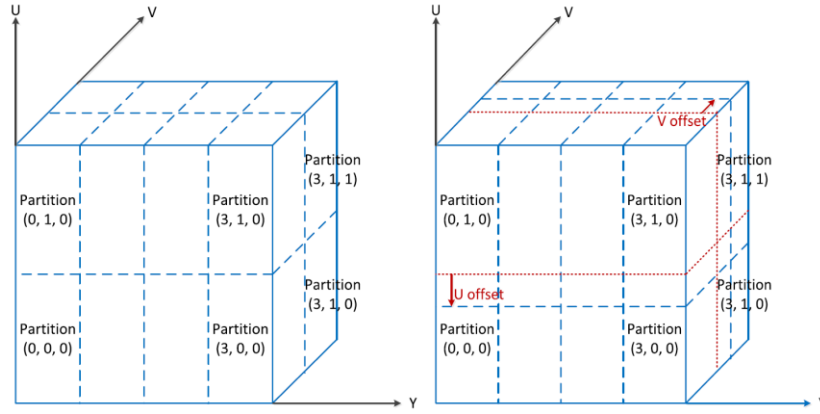


图 5 左：非对称 3D LUT；右：非均匀色度分区

不对称三维查找表是在对称 3D LUT 基础上提出的，目的是为了在保持小型 LUT 的同时实现准确的色彩映射，基底层颜色空间被分割成多个相同大小的长方体分区，其中亮度分量比两个色调分量更细，如图 5 中左图所示。对于每个长方体分区，使用以下模型的线性交叉颜色预测，其中 (Y_b, U_b, V_b) 表示基本层色域中的一个像素的三色分量，而 (Y_e, U_e, V_e) 表示相关像素在增强层色域中， a 和 b 表示映射参数。在配置文件 Scalable Main 10 中可以支持多达 $8 \times 2 \times 2$ 长方体分割，由此 LUT 项的最大量可以达到 384 ($8 \times 2 \times 2 \times 4 \times 3$)，相对于对称的 3D LUT 方法来说数量少了很多。

$$\begin{bmatrix} Y_e \\ U_e \\ V_e \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_b \\ U_b \\ V_b \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$$

非均匀速度分区是在了解到色度分量的样本值通常分布不均匀的事实，而提出的。图 5 的右侧图呈现了不均匀色度分割的例子，作为参考，均匀色度分区由虚线表示。为了定位色度分区边界，相对于虚线中的均匀划分的偏移在比特流中用于两个色度分量的信号。

相位调整是为 4:2:0 的色度格式输入视频的特别设计，在这种情况下，色度分量的空间

分辨率是在水平和垂直方向上的亮度分量的空间分辨率的一半。此外，色度样本位置通常与亮度样本位置不对齐。HEVC 标准中的 VUI 定义了六个色度位置类型以指定相对于亮度样本位置的色度样本位置。CGS 中的相位调整集中在几种最常见的情况，如图 6 所示。

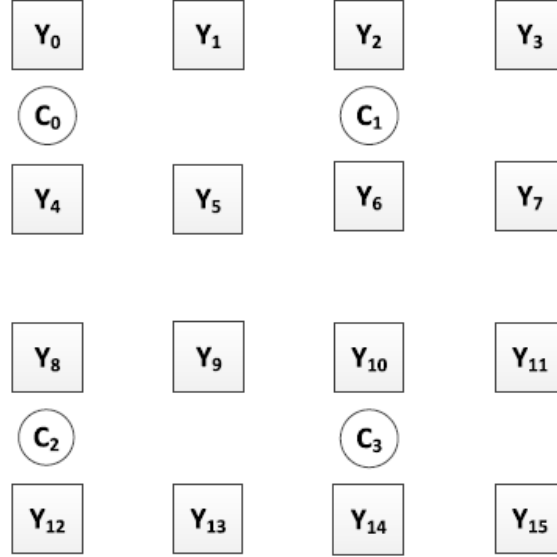


图 6 4:2:0 色度模型的常见模式

图 6 给出了一种典型的邻域亮度和色度样本的设定, 是由 BT.709 规定的 4:2:0 色度格式。然而, 在 BT.2020 中, 色度采样位置与亮度样本位置的对应关系是色度格式 4:2:0, 亮度样本位置用 Y 表示, 色度采样位置用 C 表示。为了将不对称的基于 3D LUT 的颜色映射应用于这三种颜色分量中的任意一种, 需要输入一个三元数组(Y, U, V)。为了提高颜色映射处理的精度, 在使用交叉颜色分量进行操作之前需要将不同颜色分量的样本位置对齐。具体而言, 当计算亮度分量的色彩映射输出时, 色度采样值被调整为与它们应用的对应亮度样本位置对齐。将 C 表示为输入色度样本 U 或 V 之一, Y 表示输入亮度信号, 根据 Y 所在的位置, 使用简单的 2 抽头或 4 抽头滤波器调整 C 的值, 方法如下:

$$\begin{aligned}
 c'(Y_4) &= (c(C_0) * 3 + c(C_2) + 2) / 4 \\
 c'(Y_5) &= ((c(C_0) + c(C_1)) * 3 + (c(C_2) + c(C_3)) + 4) / 8 \\
 c'(Y_8) &= (c(C_2) * 3 + c(C_0) + 2) / 4 \\
 c'(Y_9) &= ((c(C_2) + c(C_3)) * 3 + (c(C_0) + c(C_1)) + 4) / 8
 \end{aligned}$$

类似的, 当计算色度分量的颜色映射输出时, 调整亮度样本值以与它们应用的相应色度采样位置对准。也就是说, 根据 c 所在的位置, 使用简单的 2 抽头滤波器调整 y 的值, 方法如下:

$$\begin{aligned}
 y'(C_0) &= (y(Y_0) + y(Y_4) + 1) / 2 \\
 y'(C_1) &= (y(Y_2) + y(Y_6) + 1) / 2 \\
 y'(C_2) &= (y(Y_8) + y(Y_{12}) + 1) / 2 \\
 y'(C_3) &= (y(Y_{10}) + y(Y_{14}) + 1) / 2
 \end{aligned}$$

SHVC 有灵活的方式通过有效的信令来支持位深的可扩展性 (BDS), CGS 中的颜色映射过程以 CGS 参考信号输入和输出的位深, 这样也可以支持 BDS。SHVC 中的上采样过程通过调整上采样内的位移数量, 根据需要对输入和输出之间的差异进行补偿。支持 BDS 的一种方法是简单地通过相应地制定输入和输出位深来应用组合位深转换和空间分辨率转换的上采样过程, 过程很简单, 不需要考虑上采样过程的输入和输出之间的颜色特征的变化。这一方法的层间过程在图 7(a)给出, 用于 CGS、BDS 和空间组合可扩展性。第二种选择是将位深转换过程和色域转换过程组合在一起。第二种方法的层间过程由图 7(b)给出, 与第一种方法相比, 第二种方法可以保持足够的转换精度并具有更好的编码性能。

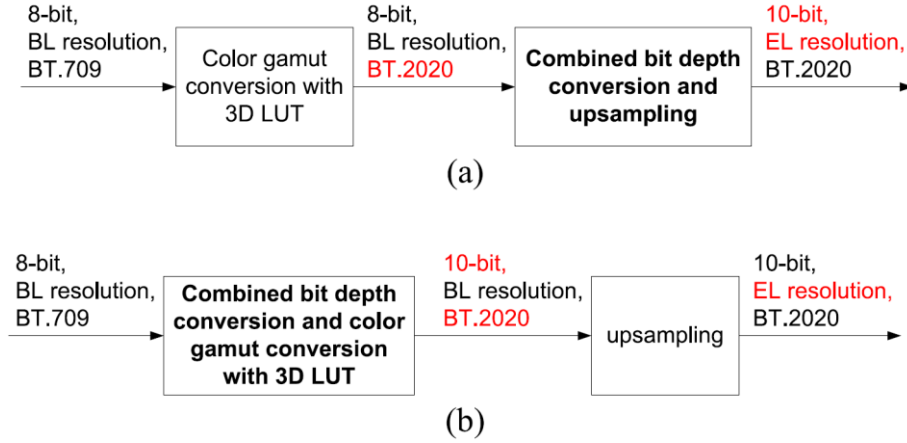


图 7 SHVC 层间处理过程

11.3 其他相关编码方案

前面内容介绍了当前 HEVC 标准中包含的 HDR/WCG 相关的编码技术。在本节中, 简要介绍其他的一些在 HEVC 标准扩展过程中提出的一些代表性的编码相关技术。

11.3.1 替换颜色空间

在传统方法中, 视频编码使用 Y'CbCr 颜色空间, 该颜色空间的目的是对 R, G 和 B 颜色分量进行去相关。在 SMPTE ST 2085 中标准化的新 YDzDx 色差空间被添加到 HEVC 版本 3 中的 VUI [1]。YDzDx 通过对 XYZ 原始信号的差分矩阵求导获得, 其方式与 Y'CbCr 从 RGB 导出的方式类似。从 XYZ 颜色空间到 YDzDx 空间的变换如下指定 (Y 分量与 XYZ 颜色空间中的 Y 一致):

$$D_Z = \frac{c_1 \times Z - Y}{2.0} \quad \text{and} \quad D_X = \frac{X - c_2 \times Y}{2.0}$$

c_1 和 c_2 是常数, 旨在改善由色度下采样和量化噪声产生的白平衡偏移。在过去十年中, 类似的 XYZ 色差衍生空间已被用于数字电影应用中。YDzDx 允许 XYZ 视频通过 4: 2: 0 设备链, 与 4: 4: 4 色度格式的 XYZ 直接编码相比, 提供一点额外的编码性能。电影制片厂 XYZ 主要空间的主要动机是提供更加面向未来的色彩坐标系统, 通常用于工作室制作工作流程,

根据定义，它完全包含人类视觉的视觉范围，从而避免单独的渲染和分级，也避免了随着时间推移颜色空间的逐步增加。

针对色调线性度优化的其他色彩空间，例如 CIELAB 和 IPT [2], [3]，近年来也在持续针对 HDR 和 WCG 图像和视频编码进行一些研究。

11.3.2 替换编码框架

已经有许多与 HDR 内容的编码有关的工作，主要基于色调映射操作以将 HDR 信号变换为 8 或 10 位整数低动态范围 (LDR) 信号。这些方法可以分为三类。

第一类是基于比特深度的感知域中的 HDR 信号的表示，其高于应用于 SDR 视频的通常的 8 比特 (参见[4], [5])。传递函数的研究属于这一类。该类工作一般使用单层解决方案对得到的转换信号进行编码。

第二类基于调制原理，并使用两个层 (参见[6], [7])。基础层是 SDR 视频，附加层包含调制信息。在解码器侧，将解码的 LDR 内容乘以解码的调制比以重建 HDR 内容。

第三类还使用两个层，其中基于逆色调映射的层间预测来预测来自 LDR 基础层的 HDR 层。预测残差被编码为增强层 (参见[8] - [10])。即使从概念上讲它们可以应用于 HEVC，这些工作最初也是基于 MPEG-2 或 AVC。

除了已经在 HEVC 中指定的技术之外，JCT-VC 也收到几个候选解决方案以用于 HDR 和 WCG 支持。这些提议目标各不相同，比如，比现有的基于 HEVC 的编码链改进编码效率，比特流与传统解码器的向后兼容性，以及解码视频与传统显示器的向后兼容性。大多数提案仅影响编码和解码过程之前和之后的预处理和后处理步骤，因此不影响 HEVC 规范的核心编码工具。然而，其中一些提案可能需要另一个配置文件，而不是 HEVC main 或 main10 配置文件。

结合较小位深度实现的高位深度编码的问题在[11] - [14]中得到解决。常见的想法是重用 HEVC 8 (或 10) 位设计来编码更高位深度的内容，而无需例如在各种变换级，量化和熵编码级中扩展内部寄存器的位深度。

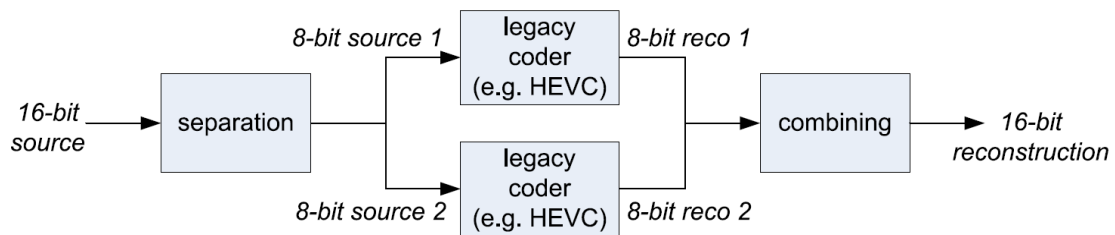


图 8 低比特深度编码器进行高比特深度信号编码

如图 8 所示，通过分割高位深度的最高有效位 (MSB) 和最低有效位 (LSB)，首先将输入 HDR 图像转换为有限位深度的两个图像 (通常每个分量 8 位) 样本。考虑到 MSB 和 LSB 之间的差异，编码过程中采用自适应量化调整。MSB 部分被编码为主图像，LSB 图像被编码为辅助图像。该工作使用 SEI 信息以在解码器处组合 MSB 和 LSB 部分。

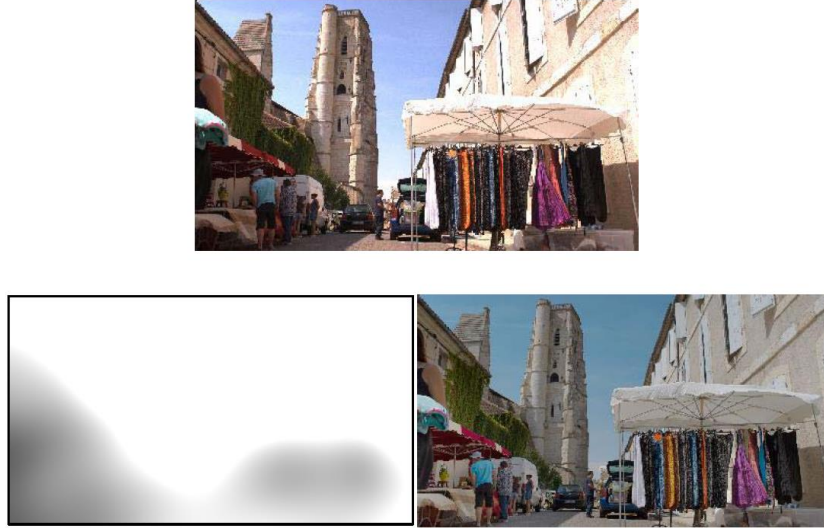


图 9 将 HDR 图像（顶部）分解为调制图像（左下）
和有限深度的 SDR 信号（右下）

基于信号调制概念的替代解决方案在[15] - [17]中提交给 JCT-VC。所提出的解决方案旨在提供改进的编码效率和 SDR 后向兼容性。原理是，即使图像全局为高位深度或 HDR，也可以在局部进行低位深度的分解。该解决方案包括编码前的预处理步骤，将输入信号分成两个独立的低位深度信号，一个调制信号（称为调制图像，为输入信号的低通版本）和另一个有限深度信号（SDR 信号）。一旦提取了低频部分，比特深度信号（SDR 图像）对应于剩余部分，如图 9 所示。

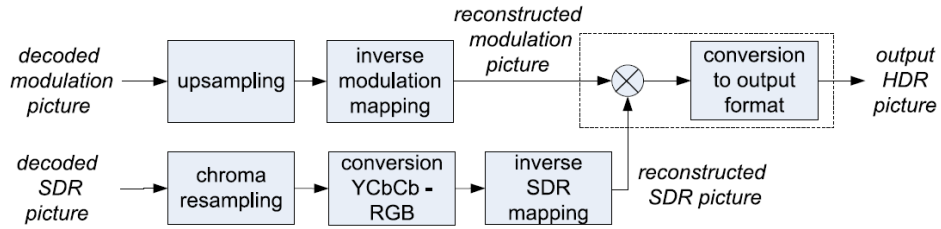


图 10 调制方法的解码器端

在解码端，相应的后处理步骤（图 10 中所示）被应用于重构的调制图像和 SDR 图像以生成输出的高位深度图像。一旦解码的 Y'CbCr 4: 2: 0 SDR 信号被解码，则将其重采样为 4: 4: 4 色度格式，然后转换为 RGB 空间。从 SDR 信号（RSDR, BSDR, GSDR）和调制信号 μ 重建 HDR 信号（RHDR, BHDR, GHDR）可以用以下简化的等式表示：

$$\begin{aligned} R_{\text{HDR}} &= f_{\text{Modul}}(\mu) \times f_{\text{SDR}}(R_{\text{SDR}}) \\ G_{\text{HDR}} &= f_{\text{Modul}}(\mu) \times f_{\text{SDR}}(G_{\text{SDR}}) \\ B_{\text{HDR}} &= f_{\text{Modul}}(\mu) \times f_{\text{SDR}}(B_{\text{SDR}}) \end{aligned}$$

在编码端，调制信号和 SDR 信号都可以（均由线性光输入 HDR 信号产生）通过两个不

同的传递函数被映射到非线性表示。必须使用逆映射函数 f_{SDR} 和 f_{Modul} 在解码器侧反转这些操作。组合两个信号的乘法应用于线性光域。完整过程独立地适用于每个像素。

为了在解码端重建 HDR 信号，改提案还提出了新的 SEI 信息（调制信道信息 SEI 信息）以嵌入从解码的调制和解码的 SDR 图像重建 HDR 信号所需的参数，该信息还指示输出重建信号的格式。

参考文献

- [1] C. Fogg, J. Helman, M. Smith, and M. Zink, Content Light Level Information SEI, document JCTVC-T0101, Geneva, Switzerland, Feb. 2015.
- [2] G. M. Johnson, X. Song, E. D. Montag, and M. D. Fairchild, “Derivation of a color space for image color difference measurement,” *Color Res. Appl.*, vol. 35, no. 6, pp. 387–400, 2010.
- [3] M. D. Fairchild and D. R. Wyble, “HDR-CIELAB and HDR-IPT: Simple models for describing the color of high-dynamic-range and wide-color gamut images,” in *Proc. IS&T/SID 18th Color Imag. Conf.*, San Antonio, TX, USA, 2010, pp. 322–326.
- [4] A. Motra and H. Thoma, “An adaptive LogLuv transform for high dynamic range video compression,” in *Proc. 17th IEEE ICIP*, Hong Kong, Sep. 2010, pp. 2061–2064.
- [5] Y. Zhang, E. Reinhard, and D. Bull, “Perception-based high dynamic range video compression with optimal bit-depth transformation,” in *Proc. 18th IEEE ICIP*, Brussels, Belgium, Sep. 2011, pp. 1321–1324.
- [6] C. Lee and C.-S. Kim, “Rate-distortion optimized compression of high dynamic range videos,” in *Proc. 16th Eur. Signal Process. Conf.*, Lausanne, Switzerland, Aug. 2008, pp. 1–5.
- [7] G. Ward and M. Simmons, “Subband encoding of high dynamic range imagery,” in *Proc. 1st Symp. Appl. Perception Graph. Visualizat.*, Aug. 2004, pp. 83–90.
- [8] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. Ward, and W. Heidrich, “Optimizing a tone curve for backward-compatible high dynamic range image and video compression,” *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1558–1571, Jun. 2011.
- [9] A. Koz and F. Dufaux, “Optimized tone mapping with perceptually uniform luminance values for backward-compatible high dynamic range video compression,” in *Proc. IEEE VCIP*, San Diego, CA, USA, Nov. 2012, pp. 1–6.
- [10] R. Mantiuk, A. Efremov, K. Myszkowski, and H.-P. Seidel, “Backward compatible high dynamic range MPEG video compression,” in *Proc. ACM SIGGRAPH*, Saarbrücken, Germany, 2006, pp. 713–723.
- [11] E. François, C. Gisquet, G. Laroche, and P. Onno, AHG18: On 16-Bits Support for Range Extensions, JCTVC-N0142, Vienna, Austria, Jul. 2013.
- [12] W.-S. Kim, W. Pu, J. Chen, Y.-K. Wang, J. Sole, and M. Karczewicz, AhG 5 and 18: High Bit-Depth Coding Using Auxiliary Picture, JCTVC-O0190, Geneva, Switzerland, Oct. 2013.
- [13] A. Aminlou and K. Ugur, On 16-Bit Coding, JCTVC-P0162, San Jose, CA, USA, Jan. 2014.
- [14] C. Auyeung and J. Xu, AhG 5 and 18: Cross-Check of Coding of High Bit-Depth Source With Lower Bit-Depth Encoders and a Continuity Mapping, JCTVC-P0173, San Jose, CA, USA, Jan. 2014.
- [15] S. Lasserre, F. Le Leannec, and E. François, High Dynamic Range Video Coding, document JCTVC-P0159, San Jose, CA, USA, Jan. 2014.
- [16] F. Le Leannec, S. Lasserre, and E. François, Modulation Channel SEI Message, JCTVC-R0139,

Sapporo, Japan, Jul. 2014.

[17] F. Le Léannec, S. Lasserre, and E. François, Usage of Modulation Channel for High Bit-Depth Signal Encoding, JCTVC-R0267, Sapporo, Japan, Jul. 2014.